

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 03:18:06 ; Search time 10.1508 Seconds  
(without alignments)  
1581.285 Million cell updates/sec

Title: US-09-807-933B-9

Perfect score: 2106

Sequence: 1 MKPTVAITSAVALSSA.....TFKEVTCPAELTTRSGCERK 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	714	33.9	376	1 GUNK_FUSOX	P45699 fusarium ox
2	710	33.7	213	1 GUN5_HUMIN	P43316 humicola in
3	592	28.1	511	1 GUNB_PSEFL	P18126 pseudomonas
4	525.5	25.0	393	1 GUN1_USTWA	P54424 ustilago ma
5	226.5	10.8	471	1 GUX2_TRIRE	P07987 trichoderma
6	221	10.5	210	1 P8SP_PORPU	P50272 porphyra pu
7	190	9.0	418	1 GUN2_TRIRE	P07982 trichoderma
8	177	8.4	475	1 SIM1_YEAST	P40472 saccharomyc
9	174.5	8.3	378	1 SLG1_YEAST	P54867 saccharomyc
10	172	8.2	438	1 GUX3_AGABI	P49075 agariusus bi
11	166.5	7.9	462	1 GUNB_FUSOX	P46236 fusarium ox
12	166.5	7.9	481	1 LORI_MOUSE	P18165 mus musculu
13	166.5	7.9	503	1 YN23_YEAST	P53832 saccharomyc
14	166	7.9	937	1 HYR1_CANAL	P46591 candida alb
15	163	7.7	600	1 SP96_DICDI	P14328 dictyosteli
16	161.5	7.7	662	1 MUC1_XENLA	Q05049 xenopus lae
17	159.5	7.6	2715	1 G156_PARPR	P13837 paramecium
18	158.5	7.5	537	1 SP70_DICDI	P15269 dictyosteli
19	158	7.5	786	1 STUB_DROME	Q05319 drosophila
20	158	7.5	1367	1 AMTH_YEAST	P08640 saccharomyc
21	156.5	7.4	2704	1 G168_PARPR	P17053 paramecium
22	153.5	7.3	385	1 GUNF_FUSOX	P46239 fusarium ox
23	153	7.3	242	1 GUN5_TRIRE	P43317 trichoderma
24	151.5	7.2	962	1 GUNA_PSEFL	P10476 pseudomonas
25	151	7.2	1246	1 YWV2_CABEL	P34504 canoxyhabdi
26	148.5	7.1	389	1 SRR1_BOMMO	P07856 bombyx mori
27	146	6.9	542	1 SCWB_YEAST	P53189 saccharomyc
28	146	6.9	605	1 YHC8_YEAST	P38739 saccharomyc
29	145.5	6.9	1150	1 APMU_PIG	P12021 sus scrofa
30	145	6.9	1218	1 JAG1_MOUSE	Q9qrx0 mus musculu
31	144.5	6.9	1408	1 SERR_DROME	P18168 drosophila
32	144	6.8	513	1 GUX1_TRIRE	P00725 trichoderma
33	144	6.8	686	1 DLL4_MOUSE	Q9ji71 mus musculu

34	144	6.8	757	1 VGLX_HSVBB	P28968 equine herp
35	142.5	6.8	555	1 DP87_DICDI	Q04503 dictyosteli
36	142.5	6.8	573	1 C114_MOUSE	P19467 mus musculu
37	142	6.7	513	1 GUX1_TRIVI	P19355 trichoderma
38	142	6.7	1218	1 JAG1_HUMAN	P78504 homo sapien
39	141.5	6.7	1200	1 ICEN_PSESY	P06620 pseudomonas
40	141.5	6.7	1224	1 PER_DROME	P07663 drosophila
41	141	6.7	1561	1 PPRB_HUMAN	Q15648 h peroxisom
42	141	6.7	2471	1 NTC2_HUMAN	Q04721 homo sapien
43	140.5	6.7	459	1 GUN1_TRIRE	P07981 trichoderma
44	140.5	6.7	540	1 GUX1_ASPAC	O59843 aspergillus
45	139.5	6.6	463	1 GUN1_TRILO	Q12714 trichoderma

ALIGNMENTS

RESULT 1  
GUNK\_FUSOX  
ID GUNK\_FUSOX STANDARD; PRT; 376 AA.  
AC P45699;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Putative endoglucanase type K precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase).  
OS Fusarium oxysporum.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Hypocreales; mitosporic Hypocreales; Fusarium.  
OX NCBI\_TaxID=5507;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95047531; PubMed=7959045;  
RA Sheppard P.O., Grant F.J., Oort P.J., Sprecher C.A., Foster D.C., Hagen F.S., Upshall A., McKnight G.L., O'Hara P.J.;  
RT "The use of conserved cellulase family-specific sequences to clone cellulase homologue cDNAs from Fusarium oxysporum.";  
RL Gene 150:163-167(1994).  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose.  
CC -!- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).  
CC -!- SIMILARITY: BELONGS TO CELLULOSE FAMILY K (FAMILY 45 OF GLYCOSYL HYDROLASES).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L29381; AAA65589.1; -.  
HSSP; P43316; 2ENG.  
InterPro; IPR000254; CBD\_fungal.  
InterPro; IPR000334; GH\_45.  
Pfam; PF00734; CBM\_1; 1.  
Pfam; PF02015; Glyco\_hydro\_45; 1.  
SMART; SM00236; fCBD; 1.  
PROSITE; PS00562; CBD\_FUNGAL; 1.  
PROSITE; PS01140; GLYCOSYL\_HYDROL\_F45; 1.  
Cellulose degradation; Hydrolase; Glycosidase; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 376 PUTATIVE ENDOGLUCANASE TYPE K.  
FT DOMAIN 19 308 CATALYTIC.  
FT DOMAIN 309 338 LINKER.  
FT DOMAIN 339 376 CELLULOSE-BINDING.  
FT ACT\_SITE 29 29 NUCLEOPHILE (BY SIMILARITY).  
FT ACT\_SITE 140 140 PROTON DONOR (BY SIMILARITY).  
SQ SEQUENCE 376 AA; 39235 MW; B430A5F962B9F882 CRC64;  
Query Match 33.9%; Score 714; DB 1; Length 376;  
Best Local Similarity 59.1%; Pred. No. 2.8e-38;







```
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative polysaccharide binding protein precursor.
OS Porphyra purpurea.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RA Liu Q., der Meer J.P., Reith M.E.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 FUNGAL-TYPE CELLULOSE-BINDING DOMAINS
CC (CBD).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; U08843; AAA61792.1; -.
CC HSSP; P00725; 2CBH.
CC InterPro; IPR000254; CBD_fungal.
CC Pfam; PF00734; CBM_1; 4.
CC SMART; SM00236; fCBD; 4.
CC PROSITE; PS00562; CBD_FUNGAL; 4.
CC Repeat; Signal.
CC SIGNAL 1 22
CC CHAIN 23 210 PUTATIVE POLYSACCHARIDE BINDING PROTEIN.
CC DOMAIN 23 62 POLYSACCHARIDE BINDING DOMAIN I.
CC DOMAIN 63 105 POLYSACCHARIDE BINDING DOMAIN II.
CC DOMAIN 125 165 POLYSACCHARIDE BINDING DOMAIN III.
CC DOMAIN 166 210 POLYSACCHARIDE BINDING DOMAIN IV.
CC SEQUENCE 210 AA; 21967 MW; 83782D98B6863C8 CRC64;
CC
CC Query Match 10.5%; Score 221; DB 1; Length 210;
CC Best Local Similarity 29.1%; Pred. No. 1.4e-07;
CC Matches 69; Conservative 22; Mismatches 74; Indels 72; Gaps 11;
CC
CC 11 AVALALSSAAAS-CSSVYQCGIGWGTGTCDDAGSTCKAQDKNKYSQC--IPKPKG 67
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 8 AAALTLSAAASACGVLYEQCGIGDPGVTCCSEGLWC--MKMGPIYSQCRAMPGMWG 65
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 68 SSSSSSCSVSYQCGIGWGTGTCESGSTCVAQEGNKYSQCLPGSHNNAGNASSTKK 127
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 66 QVKE-----YQCGGMNYSKWTWCSPGFKCV--ELNEFFSQ-----DLANKSPV 108
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 128 TSTKTSTTAKATATVTKTKTKTKTKTKTTTAAASTSTSSAGYKVISGGKSGSGS 187
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 109 ATPKVPSTSPGPAQVCKEY-----AACGGEMFNGAK----- 141
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 188 TTRYWDCC-----ASCSWPGKASVTPVDTCASNGISLLDNAQSGCGNGGFM 237
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 142 -----CKGFLGVCYETSGKWSQCRAPP-----KMGVCRYAQCGGNGYM 182
CC |||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC
CC RESULT 7
CC GUN2 TRIE
CC ID_GUN2 TRIE STANDARD; PRT; 418 AA.
CC AC P07982;
CC DT 01-AUG-1988 (Rel. 08, Created)
CC DT 01-AUG-1988 (Rel. 08, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE Endoglucanase EG-II precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase)
CC (Cellulase).
CC GN EGII OR EGIII.
CC OS Trichoderma reesei (Hypocrea jecorina).
CC OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
CC OC Hypocreales; Hypocreaceae; Hypocrea.
CC OX NCBI_TaxID=51453;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VTT-D-80133;
RX MEDLINE=98255850; PubMed=3384334;
RA Saloheimo M., Lehtovaara P., Penttilae M., Teeri T.T., Staahlberg J.,
RA Johansson G., Pettersson G., Claysons M., Tomme P., Knowles J.K.C.;
RT "EGIII, a new endoglucanase from Trichoderma reesei: the
RL Gene 63:11-21(1988).
RN [2]
RP ACTIVE SITE GLU-350.
RX MEDLINE=93131031; PubMed=8093602;
RA Macarron R., van Beumen J., Henriessat B., de la Mata I.,
RA Claeysens M.;
RT "Identification of an essential glutamate residue in the active site
RL of endoglucanase III from Trichoderma reesei.";
CC -1- FUNCTION: THE BIOLOGICAL CONVERSION OF CELLULOSE TO GLUCOSE
CC GENERALLY REQUIRES THREE TYPES OF HYDROLYTIC ENZYMES:
CC (1) ENDOGLUCANASES WHICH CUT INTERNAL BETA-1,4-GLUCOSIDIC BONDS;
CC (2) EXOCELLULOBIODOLASES THAT CUT THE DISSACCHARIDE CELLOBIOSE
CC FROM THE NONREDUCING END OF THE CELLULOSE POLYMER CHAIN;
CC (3) BETA-1,4-GLUCOSIDASES WHICH HYDROLYZE THE CELLOBIOSE AND OTHER
CC SHORT CELLO-OLIGOSACCHARIDES TO GLUCOSE.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
CC linkages in cellulose.
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC -1- CAUTION: WAS ORIGINALLY CALLED ENDOGLUCANASE EG-III.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M19373; AAA34213.1; -.
CC PIR; S28372; S28372.
CC HSSP; P00725; 2CBH.
CC InterPro; IPR000254; CBD_fungal.
CC InterPro; IPR001547; GH_5.
CC Pfam; PF00150; cellulase; 1.
CC Pfam; PF00734; CBM_1; 1.
CC ProDom; PD001821; CBD_fungal; 1.
CC SMART; SM00236; fCBD; 1.
CC PROSITE; PS00562; CBD_FUNGAL; 1.
CC PROSITE; PS00659; GLYCOSYL HYDROL F5; 1.
CC Cellulose degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
CC SIGNAL 1 21
CC CHAIN 22 418 ENDOGLUCANASE EG-II.
CC DOMAIN 22 57 CELLULOSE-BINDING (BY SIMILARITY).
CC LINKER.
CC DOMAIN 58 91 CATALYTIC.
CC DOMAIN 92 418 PYRROLIDONE CARBOXYLIC ACID
CC MOD RES 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 124 124 BY SIMILARITY.
CC DISULFID 29 46 BY SIMILARITY.
CC DISULFID 40 56 BY SIMILARITY.
CC ACT SITE 239 239 PROTON DONOR (BY SIMILARITY).
CC ACT SITE 350 350 NUCLEOPHILE.
CC SEQUENCE 418 AA; 44227 MW; 26A492D55237A49B CRC64;
CC
CC Query Match 9.0%; Score 190; DB 1; Length 418;
CC Best Local Similarity 24.3%; Pred. No. 2.2e-05;
CC Matches 83; Conservative 41; Mismatches 119; Indels 98; Gaps 15;
CC
CC 76 SVYSCCGIGWGTGTCESGSTCVAQEGNKYSQCLPGSHNNAGNASSTKTKTSTT 135
CC :|||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 24 TWVQCGIGWGTGTCESGSTCVAQEGNKYSQCLPGSHNNAGNASSTKTKTSTT 60
CC :|||:|||||:|||||:|||||:|||||:|||||:|||||:
CC 136 TAKATATVTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 195
CC :|||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 61 -----TTTITTPRSGGTTTTRATSTSSPPSSGVRPAGVNIAGDFGCTTD--GTC 113  
 Qy 196 KASCSWPKASVGTVPDTCASN-----GSLLDANAQSCNGCGMCMNNQFMAVNDLA 251  
 Db 114 VTSKYVPELKNFTG-----SNNYPDGIQOMQHFVNE--DGMTIFRLPVGMQVLYVNNLIG 165  
 Qy 252 YGFAAASAGSNEAGWCCGCVELTFTSAAASKKMWVQVTN-----TGGDLGSHNFDLQMP 307  
 Db 166 GNLSTSTISKYDQL--VQGCLSL-----GAYCIVDIHNYRWNGGIIGQ----- 207  
 Qy 308 GGG-----VGIFNGCAQWGA-----PND-----GWGARYGGV-----SSVS 339  
 Db 208 -GGPTNAQFTSLMSQLASKYASQSRVFGWIFGNEPHDVNINWTAATVQVWTAIRNAGATS 266  
 Qy 340 DCASLPALQACQKRWFRNFKNSDNPMTFTFKEVTCPAELTT 380  
 Db 267 QFISLP-----GNDWQAGAFISDGSAAALSQVTPNPDGSTT 302

## RESULT 8

SI1 YEAST  
 ID SI1 YEAST STANDARD; PRT; 475 AA.  
 AC P40472;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE SI1 protein precursor  
 GN SI1 OR PBP3 OR YIL123W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis B., Lyne G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP IDENTIFICATION.  
 RX MEDLINE=9615611; PubMed=8574583;  
 RA Dahmann C., Diffley J.F.X., Nasmyth K.A.;  
 RT "S-phase-promoting cyclin-dependent kinases prevent re-replication by  
 RT inhibiting the transition of replication origins to a pre-replicative  
 RT state.";  
 RL Curr. Biol. 5:1257-1269 (1995).  
 CC -!- FUNCTION: REQUIRED FOR THE MAINTENANCE OF THE CLB5 KINASE  
 CC ACTIVITY.  
 CC -!- SIMILARITY: BELONGS TO THE SUN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; 246833; CAA86869.1; -;  
 DR SGD; S0001385; SI1.  
 KW SIGNAL.  
 FT SIGNAL. 1 19 POTENTIAL.  
 FT CHAIN 20 475 SI1 PROTEIN.  
 FT DOMAIN 58 112 ALA-RICH.  
 FT DOMAIN 80 203 SER-RICH.  
 FT CARBOHYD 422 422 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 475 AA; 48070 MW; B7B0F0BC8478612F CRC64;

Query Match 8.4%; Score 177; DB 1; Length 475;  
 Best Local Similarity 24.6%; Pred. No. 0.00016;  
 Matches 98; Conservative 51; Mismatches 139; Indels 110; Gaps 18;  
 Qy 4 TVAITSI AVALSSAAASCSVYQCGIGWGTGPTCCDAGSTCKAQDKNY----YS 59  
 Db 67 TSAIITTSV-LAPTSSAAAGIAASI-----AVSSAALAKNEKISDRAAS 109  
 Qy 60 QCIPKPGSSSSSSCSVYQCGIGWGTGPTCCBSSGTCVAQEGNKYISQCLPQSHSNA 119  
 Db 110 ATASTSQASSSSSSSSSATS-----TLESSTV-----SSSSSEA 143  
 Qy 120 GNASSTKTKSTTTAKAT--ATVTKTVTKTTTKTSTTAAASTSTSSAGYKV 177  
 Db 144 APTSTVTVSTSATOSSASSATKSTSTSTSTSTSTSTSTSTSTSTSTSTSTSSSSSSSSSSSSGGS 203  
 Qy 178 ISGGKSG-SGSTTRYWD---CCK-----ASCWPKASVGTVPDTCASNGISLDAN 225  
 Db 204 IYGLADFSGPSEKFDQGTIPCDKFPSSQGVISIDWIEGGWGSVENTDTSTGSCKEGS 263  
 Qy 226 AQS-GCNGG-----NGFMNNOQPAWVNDLAY-----GFAAASIAAG-- 261  
 Db 264 YCSYSCOPGMSKTOWPSDQPSDGRSVGGLCKNGLYRSNTDADYLCWGVAAVYVSKL 323  
 Qy 262 SNEAGWCCGYE-----LTFTS GAASGKMWVQ-----VTNTGDLGSHNFDLQMPGG 310  
 Db 324 SKGVAICRTDYPGTENMVIPTVEGSSSLPTVVDDQDTYFTWEKTKTSAQYI---VNNAG 380  
 Qy 311 VGIFNGCAQWGAAPNDGNG-----ARYGGVSSVS 339  
 Db 381 VSVEDGCI--WGTSGSIGNWAPLNFAGAGTGGVTVLS 416  
 RESULT 9  
 SI1 YEAST  
 ID SI1 YEAST STANDARD; PRT; 378 AA.  
 AC P54857;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE SI1 protein precursor  
 GN SI1 OR YOR008C OR UNF378.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA de Bettignies G., Bergez-Aullo P., Barthe C., Louvet O.,  
 RA Peypouquet M.F., Morel C., Doignon F., Crouzet M.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051599; PubMed=8896276;  
 RA Sterky F., Holmberg A., Pettersson B., Uhlen M.;  
 RT "The sequence of a 30 kb fragment on the left arm of chromosome XV  
 RT from Saccharomyces cerevisiae reveals 15 open reading frames, five of  
 RT which correspond to previously identified genes.";  
 RL Yeast 12:1091-1095 (1996).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U39481; AAA85862.1; -;  
 DR EMBL; U43491; AAC49488.1; -;  
 DR EMBL; Z74916; CAA99196.1; -;  
 DR SGD; S0005534; SI1.  
 DR InterPro; IPR002889; WSC.



- 1 -

RP SEQUENCE FROM N.A.  
RX MEDLINE=95047531; PubMed=7959045;  
RA Sheppard P.O., Grant F.J., Oort P.J., Sprecher C.A., Foster D.C.,  
RA Hagen F.S., Upshall A., McKnight G.L., O'Hara P.J.;  
RT "The use of conserved cellulase family-specific sequences to clone  
RT cellulase homologue cDNAs from Fusarium oxysporum.";  
RL Gene 150:163-167(1994).  
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic  
CC linkages in cellulose.  
CC -1- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).  
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY B (FAMILY 6 OF GLYCOSYL  
CC HYDROLASES).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L29377; AAA65595.1; -  
CC HSP; P07987; ICB2.  
CC InterPro; IPR000254; CBD\_fungal.  
CC InterPro; IPR001524; GH\_6.  
CC Pfam; PF00734; CBM\_1; 1.  
CC Pfam; PF01341; Glyco\_hydro\_6; 1.  
CC PRINTS; PR00733; GLYDRLAS6.  
CC ProDom; PD001821; CBD\_fungal; 1.  
CC ProDom; PD003733; GH\_6; 1.  
CC SMART; SM00236; fcbd; 1.  
CC PROSITE; PS00562; CBD\_FUNGAL; 1.  
CC PROSITE; PS00655; GLYCOSYL\_HYDROL\_F6\_1; 1.  
CC PROSITE; PS00656; GLYCOSYL\_HYDROL\_F6\_2; 1.  
CC Cellulose degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.  
KW SIGNAL 1 16  
FT CHAIN 17 462 PUTATIVE ENDOGLUCANASE TYPE B.  
FT DOMAIN 17 65 CELLULOSE-BINDING.  
FT DOMAIN 66 99 LINKER.  
FT DOMAIN 100 462 CATALYTIC.  
FT ACT\_SITE 190 190 BY SIMILARITY.  
FT ACT\_SITE 236 236 PROTON DONOR (BY SIMILARITY).  
FT ACT\_SITE 416 416 NUCLEOPHILE (BY SIMILARITY).  
FT DISULFID 33 50 BY SIMILARITY.  
FT DISULFID 44 60 BY SIMILARITY.  
FT DISULFID 191 250 BY SIMILARITY.  
FT DISULFID 383 430 BY SIMILARITY.  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 462 AA; 49207 MW; E25B2F5B828B637F CRC64;  
  
Query Match 7.9%; Score 166.5; DB 1; Length 462;  
Best Local Similarity 44.3%; Pred. No. 0.00071;  
Matches 43; Conservative 9; Mismatches 26; Indels 19; Gaps 4;  
  
QY 73 SCSS-VYSCGGIGWGGTCCSGSTCVAGQKNKYSQCLPGS--HSNAGNASTTKTS 129  
Db 24 SCNGVWACGGQWNSGTTCCTSGNKC--KLNDFYSQCQPGSAEPTSSTAAGPSST---- 77  
  
QY 130 TKTSTTTAKATATVTTTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 166  
Db 78 -----TATKTTATGSSSTTAGSVTSAPPAAS 104  
  
RESULT 12  
LORI MOUSE STANDARD; PRT; 481 AA.  
AC P18165;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
  
DT DE 15-JUN-2002 (Rel. 41, Last annotation update)  
GS Loricrin.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90275605; PubMed=2190691;  
RA Mehrel T., Hohl D., Rothnagel J.A., Longley M.A., Bundman D.,  
RA Cheng C., Lichi U., Bisher M.E., Steven A.C., Steinart P.M.,  
RA Yuspa S.H., Roop D.R.;  
RT "Identification of a major keratinocyte cell envelope protein,  
RT loricrin.";  
RL Cell 61:1103-1112(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=BALB/c;  
RX MEDLINE=95256248; PubMed=7738016;  
RA Disepio D., Jones A., Longley M.A., Bundman D., Rothnagel J.A.,  
RA Roop D.R.;  
RT "The proximal promoter of the mouse loricrin gene contains a  
RT functional AP-1 element and directs keratinocyte-specific but not  
RT differentiation-specific expression.";  
RL J. Biol. Chem. 270:10792-10799(1995).  
CC -1- FUNCTION: MAJOR KERATINOCYTE CELL ENVELOPE PROTEIN.  
CC -1- SUBUNIT: MONOMERS ARE CROSSLINKED BY DISULFIDE AND N-(GAMMA-  
CC GLUTAMYL) LYSINE ISODIPEPTIDE BONDS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M34398; AAA39444.1; -  
CC EMBL; U09189; AAA82152.1; -  
CC PIR; A35628; A35628.  
CC HSP; P02876; 9WGA.  
CC MGD; MG1:96816; Lor.  
KW Keratinization.  
SQ SEQUENCE 481 AA; 37830 MW; 97349A786FF239FE CRC64;  
  
Query Match 7.9%; Score 166.5; DB 1; Length 481;  
Best Local Similarity 24.0%; Pred. No. 0.00074;  
Matches 83; Conservative 19; Mismatches 121; Indels 123; Gaps 16;  
  
QY 19 SAEAAACSSVY-----CQCGGIG---WTGPTCCDAGSTCKAKDNKYSSCIPKPKGSSS 70  
Db 222 SSGSSCGGGYSGGGSSCGGGYSGGGTSCGGSSGGGGSSGQYQCYGGGSSG 281  
  
QY 71 SSSCSSVY-----SQCQG--IGWSGPTC-----CESGSTCVAQGNKY--SQCLPESH 115  
Db 282 GSSCGGYSGGGSSCGGYSGGGSSCGGGSSCGGGSSCGGGSSCGGGSSCGGGSS 341  
  
QY 116 SNNAGNASTTKTSKTSKTSKTSKTSKTSKTSKTSKTSKTSKTSKTSKTSKTSKTS 175  
Db 342 GGGGGYSSQ-----TSQTSQAPQ 361  
  
QY 176 KVISGGKSGSTTRYWDCKKASCWPKASVTGPVDTCASNGISLLDANAQSCNGNG 235  
Db 362 QSYGGSSGGG-----GSCG--GGSSGGGGGGCYSSG-----GGSSGGGGGG-- 403  
  
QY 236 FWCNNPQWAVNDELAYGFAAASIAGSNEAGWCCGCYELTFTSAAAGKMWVQVNTG 295  
Db 404 -----YSGGGGCGGGSSGGSGGC-----GGSSG-----GSGG 433  
  
QY 296 DLGNSHFDLPQGGVGFNCAQWGAQWGAQWGAQWGAQWGAQWGAQWGAQWGAQW 341  
Db 434 GCGGGY-----SGGG-----GGSSCGGGSSGGSS--GGKGPVVC 467



